Together ahead. RUAG

EXERCISE REPORT

LOCKED SHIELDS: NATO CYBER DEFENCE EXERCISE 2012

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Schweizer Armee
Armée suisse
Esercito svizzero
Swiss Armed Forces

CDX12 Findings and Conclusions by RUAG

Florian Schütz, Director, Cyber Defence Centre, RUAG
Stefan Burschka, Head of R&D, Cyber Defence Centre, RUAG

November 12, 2012

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
08 JUL 2013	**N/A**	

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Multinational Experiment 7: Cyber 3.4 MNE7 LoE Situational Awareness Tools RUAG	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **JOINT STAFF-MN//ACT Integration 116 Lakeview Parkway Suffolk, VA 23435**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited.

13. SUPPLEMENTARY NOTES
The original document contains color images.

14. ABSTRACT
NATO Cyber Defense excercise aimed at bringing different countries together to improve international coopertion in cibyre defence.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**UU**	**13**	

Background

In March 2012, NATO's Cooperative Cyber Defence Centre of Excellence organized its 2012 Cyber Defence Exercise (CDX12), Locked Shields, aimed at bringing different countries together to improve international cooperation in cyber defence. Computer specialists from various nations gathered to improve cyber defence capabilities; more than 200 participants, based in several European countries, participated in the training.

The exercise involved Red Teams attacking Blue Teams, while Blue Teams endeavoured to defend their networks and to preserve services to customers; in parallel, White and Green Teams tried to maintain Situational Awareness, and provided technical support. The nerve centre of the exercise was at NATO's premises in Tallinn, Estonia.

As a leading defence contractor, and pioneer in cyber defence technologies, RUAG participated in the exercise by leveraging its Artificial Intelligence (AI) enabled tool, Tranalyzer; furthermore, it utilized a team of specially-trained, cutting-edge IT security and forensics specialists from its Cyber Defence Centre in Bern, Switzerland. Its goal was to verify and to enhance its cyber defence capabilities in a multilateral, real-world setting.

Through this exercise, RUAG also attempted to corroborate and ascertain the key postulates that underlie its Cyber Defence strategy, innovation, and technological investments. Namely, that Situational Awareness, i.e. the capability to precisely understand and interpret the combat theatre with a high degree of granularity, is by far the greatest challenge – and missing link – in Cyber Defence nowadays.

Challenge

To any army, one of the biggest challenges facing any modern army is being able to rely on information and communication systems that are safe, and free of threats like eavesdropping, information leakage, intrusion, and denial of service – or other kinds of cyber-menaces.

This challenge is made more pressing by the wide-spread availability of cyber-weapons, which are distributed without control. No technological, financial or legal barrier can effectively counter this availability. Furthermore, conflict in the cyberspace is asymmetrically balanced in favour of the assailant – an attack can therefore leverage the latest technological advances, while a defending party must constantly strive to maintain its defenses. Adding to the challenge, protection schemes are often inappropriate as they involve heavy operational constraints and prevent communication with "external" (i.e. non-compatible) counterparts – without guaranteeing the required level of safety.

The stake is simply put: armies must counter cyber-threats by eliminating – or inverting – the asymmetry, and effectively countering ever-changing threat landscapes. If one can reinstate symmetry in cyber-warfare, one is able to prevail through other, more conventional means; one can defend one's information pipes and thus protect one's decision processes, while damaging that of the opponents.

Technologically speaking, in the modern context, this means being able to leverage new, intuitive and innovative solutions that help deliver optimal Situational Awareness. Information networks and systems are vulnerable mostly because defence capabilities are still largely blind: walls have been erected, their strength and height have largely been debated and improved, but few watchtowers – if any – have yet been conceived or built.

When it comes to cyber-warfare, the world is still in a pre-1940 situation – where no Radar station or satellite network is available to determine the strength or direction of an enemy attack. This pervasive necessity is responsible for most vulnerabilities in modern systems.

Approach

In this exercise, RUAG leveraged its Artificial Intelligence (AI) enabled platform, Tranalyzer – as well as a team of specialists operating it. RUAG's goal was to assess the ability of Tranalyzer and its team to deliver Situational Awareness (SA), i.e. theatre understanding and visibility, in a simulated cyber-combat operation.

Such theatre intelligence is crucial in that it strongly underlies all the strategic and tactical cyberwarfare decision processes. Without SA, residual cyber-risk is unmanageable, defences cannot be organized, offensives cannot be targeted, and damage cannot be assessed. To put it another way, paraphrasing , SA transforms unknown unknowns into known unknowns and known knowns in the cyber-world, and therefore constitutes an essential component in any modern cyber defense strategy.

Just as informed decisions depend upon Situational Awareness, SA itself depends on underlying quality observations, or as they are called in the cyber-world, “quality data”. The following diagram depicts this trilateral, cyclic dependency.

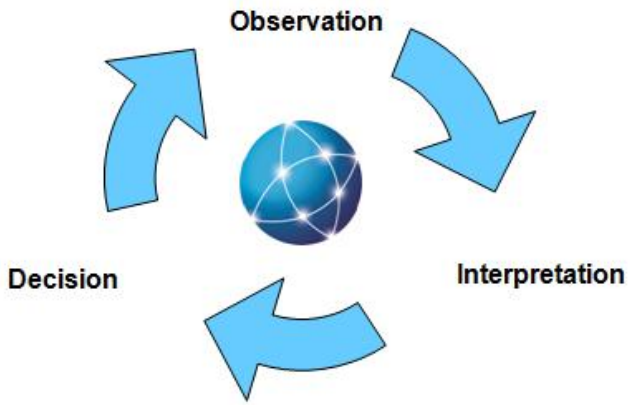

Fig 1 – Situational Awareness, i.e. the appropriate processing and interpretation of observation data, is essential for human beings to make informed strategic and tactical decisions.

A far as the observation data is concerned, anything that happens on computers or networks, down to the lowest granularity of functioning, can potentially be observed and logged; it is therefore technically and theoretically possible, unlike in the physical world, to cover one’s cyber-territory with the cyber equivalent of ubiquitous and infinite-resolution surveillance cameras and microphones.

For this reason, there exists a critical capability that one must possess in order to enjoy relevant Situation Awareness: that is, a capacity to interpret data, i.e. make sense of the massive number of events that one can collect and centralize.

One must be able to transform useless data into valuable information; it is a necessary condition for Situational Awareness, and therefore for effective defence in the cyber-world.

Ultimately of course, the information is interpreted by human beings, who are in charge of making related decisions. In that regard, visualization, reporting and communication platforms – which exist is numerous shapes and forms – are useful as they facilitate role distribution and decision processes. However, because no set of human beings – however large or qualified – can ever process the trillions of events that typically come out of large IT networks, it is necessary to rely on technology for most of the interpretation work. Without the appropriate software, observation data is a valueless commodity, and the cyber-theatres remain just as opaque.

RUAG's approach to this exercise was to observe and train its capacity at providing relevant Situational Awareness; it did this by operating its Tranalyzer platform. The platform processed billions of network events, coming from a systemic set of sensors, and tried to provide theatre intelligence useful to enable quantitatively superior decisions.

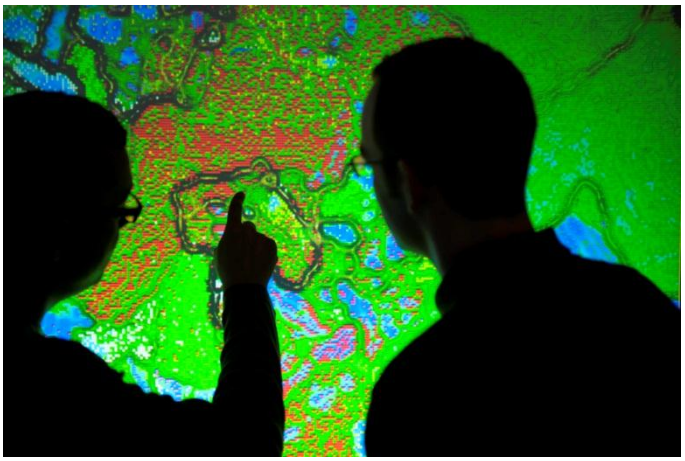

Fig 2 – During the CDX12, RUAG's aim was to assess its capacity to provide clear and relevant Situational Awareness by leveraging unique data interpretation technology.

Because RUAG's cyber-defence technology was developed in close cooperation with the Swiss Armed Forces, it was expected that Tranalyzer would perform extremely well in this cyber-warfare exercise. The assumption was that, thanks to this cooperation in particular, it would meet the needs of defence organizations, in that it would prove its relevance and effectiveness at enhancing cyber-defence operations by providing cutting-edge Situational Awareness (and thus enabling optimal decision-making).

Technical Solution

RUAG carried out the exercise by feeding the comprehensively-recorded network data of all Red and Blue teams into Tranalyzer2. The software, that leverages unsupervised trained Artificial Intelligence by using Emergent Self Organizing Maps (ESOMs), tried to make sense of that massive load – and to provide a clear Situational Awareness picture.

Tranalyzer was configured to provide the services of an Intrusion Detection System (IDS) as well as an Anomaly Detection System, under minimalistic training conditions, and with proper functioning in encrypted environments. The de-facto IDS standard (Snort) was used both as a labelling agent and as benchmark. The following two diagrams illustrate these facts.

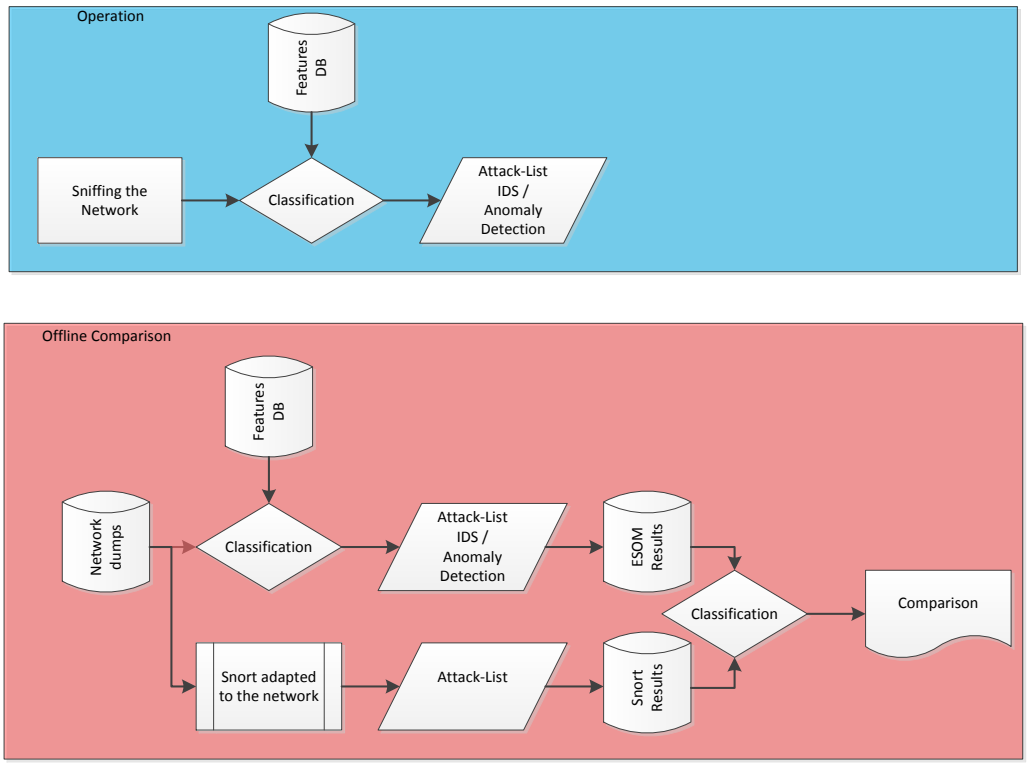

Fig 3 – In the exercise, Tranalyzer's efficiency at providing Situational Awareness was tested. It was configured as an IDS, and compared with the IDS benchmark Snort.

In the exercise, Tranalyzer2 was used in default configuration, and compiled with additional proprietary plug-ins:

- openDPI: plugin to help with the deep inspection of traffic
- regexp_snort_pcre: regular expression engine, to provide Snort labelling
- httpSniffer: Web protocol (HTTP) content analysis plugin

The ESOM implementation that was used is a special type of unsupervised trained Kohonen-type network for large neuron clusters in high-speed environments. The ESOM requires Tranalyzer's flows as input, and is able to provide a visual classification of unknown traffic in real time, even if encrypted.

Emergent Self Organizing Maps (ESOMs) such as the one shown below are a key output of the software RUAG used during this exercise, to assess and improve its capacity to deliver Situational Awareness of a cyberwarfare theatre.

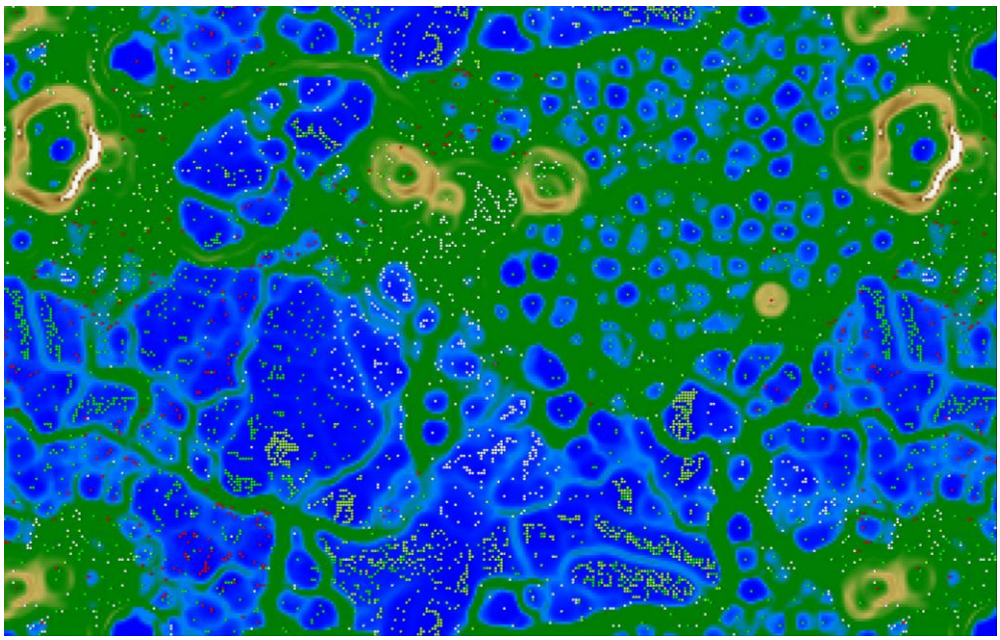

Fig 4 – An example of an ESOM map produced by RUAG's software to deliver relevant Situational Awareness.

The map can be interpreted as a landscape where blue areas denote a lower terrain and higher regions are described by green, brown or even white background (the latter denotes mountains with snow). Using Tranalyzer's proprietary AI, observations coming from the network are categorized according to various statistical similarities, and patterns (corresponding to given activities on the network) are represented as clusters – i.e. points on the map that are well grouped together, and separated from the rest by "mountains".

The lower the terrain on which dots on the map (datasets) reside, and the higher the mountains around them, the better-defined the cluster. In the example above, there are typically few such good clusters, thus illustrating a poor-quality data source.

In this next example, all regions are well separated by high walls into clear distinguishable clusters. Traffic classification accuracy without any post processing resided well above 95% in this example. Datasets are not always as good as the ones used for the map depicted below, unfortunately; ultimately, and as stated before, quality observations are essential for optimal classification, and thus for optimal intelligence.

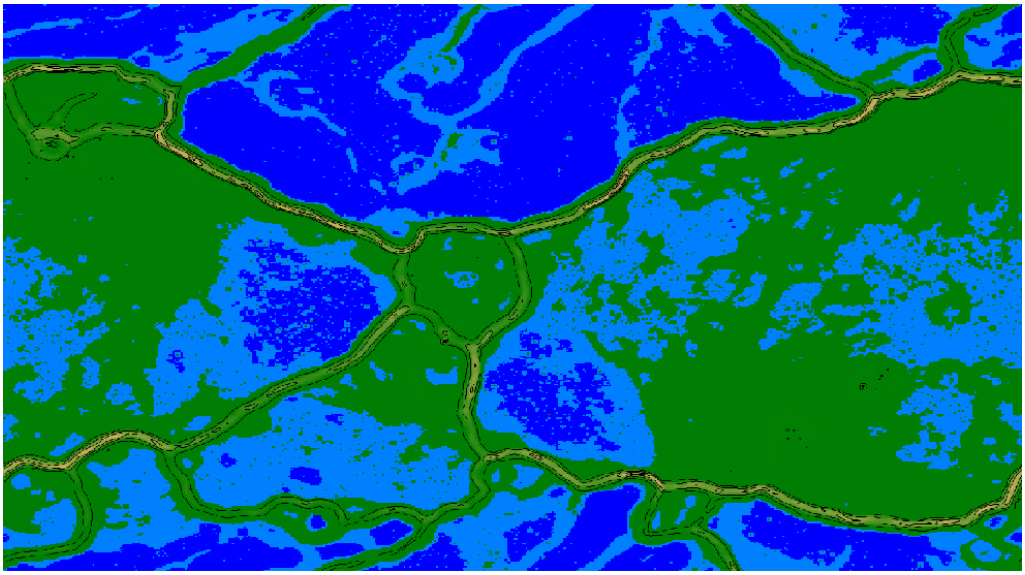

Fig 5 – An example of ESOM map where datasets are clearly clustered into well-defined regions.

Before the start of the CDX12 exercise, and after appropriate training of the software – whereby the AI is fed training data and for which interesting regions are manually defined as particular classes –the following result was obtained.

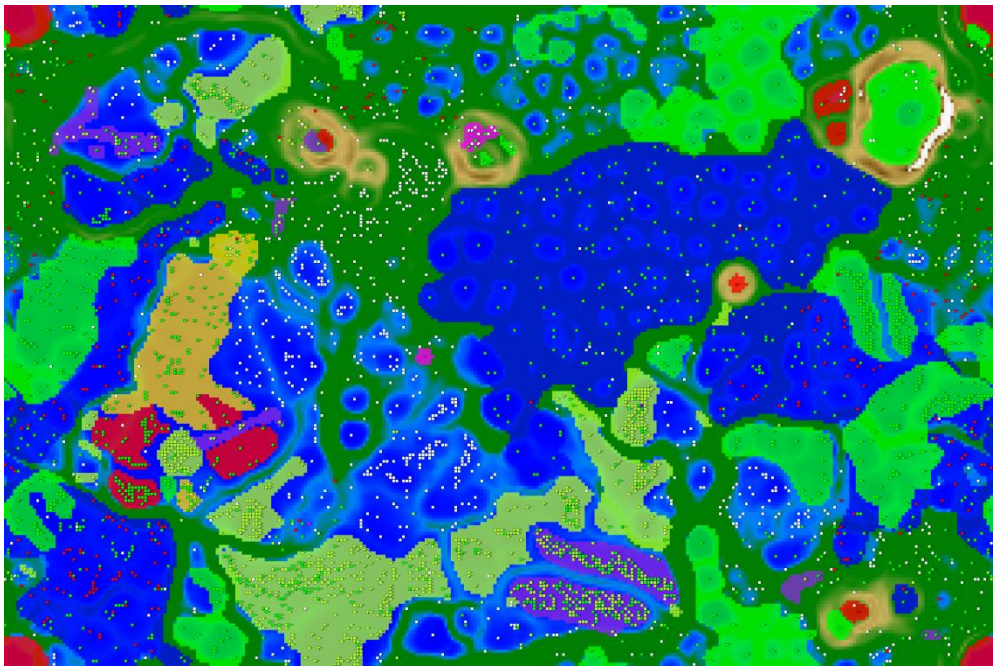

Fig 6 – The ESOM that resulted from RUAG's specialized training and fine-tuning, in which interesting behaviours have been manually pre-selected.

The following table lists the classes of activity that Tranalyzer was trained to recognize in the context of the CDX12 exercise.

Alarm	Description	Severity
00 / rest	Nothing suspicious	0
01 / green	Scan Activity	2
02 / light blue	Scan Activity with reply	2
03 / yellow	Scan Activity with or without reply	2
04 / green	Maybe Scan	1
12 / purple	Attempted user authentication	2
18 / orange	Trojan Activity	3
118 / orange	Trojan activity (high probability.) or Scan (low probability)	3
119 / orange	Trojan activity or Scan (equal probability)	3
120 / orange	Trojan activity or Scan with reply or User-login attempts (equal probability)	3
130 / red	User-login attempts or scan with reply (equal probability)	3
131 / red	User-login attempts (high probability) or scan with reply (low probability)	3

Fig 7 – Classes of alarms, as trained by RUAG for the CDX12 exercise

Although trained in this exercise for a relatively limited number of alarm classes, one can expand the ESOM's knowledge of trained features, to identify for example unknown attacks or other abnormalities. Because Artificial Intelligence is used, rather than any kind of deterministic signature recognition system, the ESOM is able to operate on encrypted or clear text traffic, indifferently. For the same reason, its ability to adapt to changing environments and threat landscapes is limitless.

Exercise Operations

The following diagram depicts the network setup used for the exercise. The ESOM classifier was positioned at the level of the firewall for all Blue Teams. While the Blue Teams were undergoing attacks from the Red Teams, the ESOM classified on-the-fly all incoming and outgoing traffic, and Tranalyzer displayed appropriate alarms on an operational screen of the War Room in Bern, Switzerland.

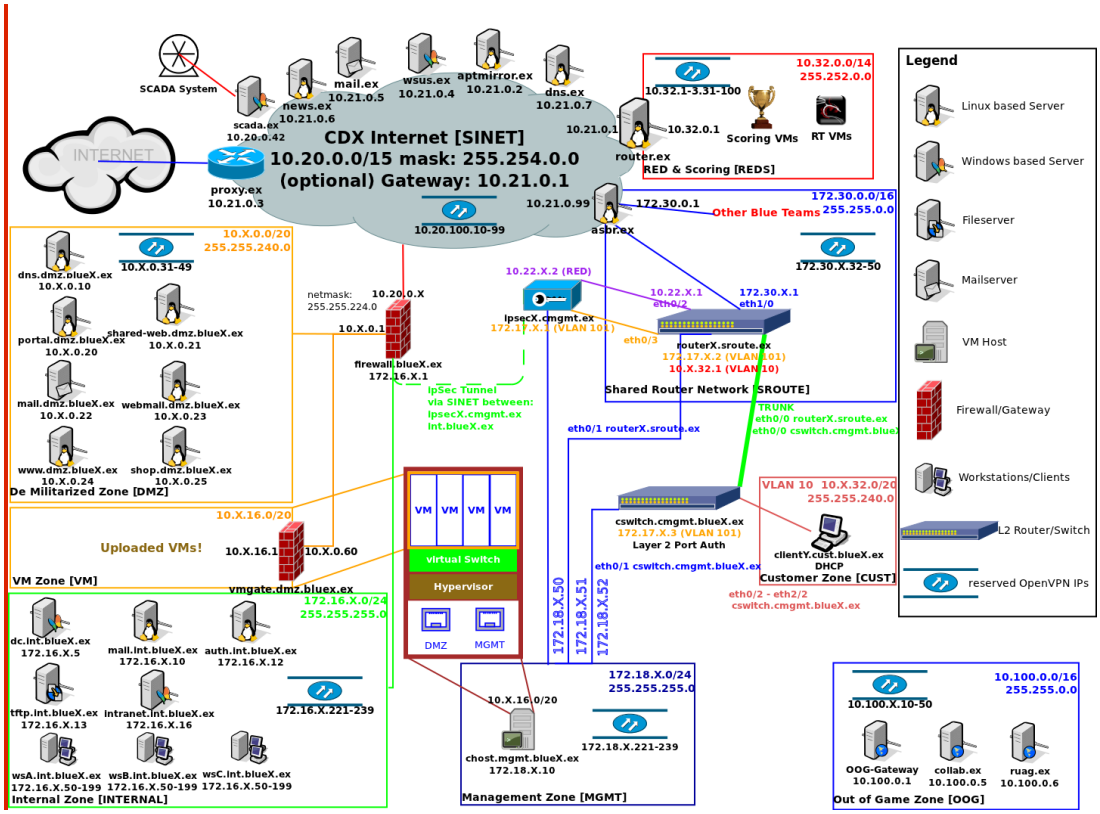

Fig 8 – The ESOM classifier was positioned at the level of the firewall used by all Blue Teams. While the attacks were occurring, it sent the appropriate alarms to a War Room in Bern, Switzerland.

The benchmark was conducted by using six Snort 2.9 sensors, using network-adapted rules (mid-March 2012). A sensor for each network was installed, e.g. DNS, Shared Zone, Customer Zone, Internal Zone, Management Zone, or Cmanagement (the Switching Network that segregated the traffic into specific IP Ranges.)

Results

Data provided by the Red Teams contained information about attacks' date/time and originating/target IP/port – and could thus be used as ground truth for the detection output of Snort (the DPI benchmark) and RUAG's Situational Awareness software (Tranalyzer/ESOM) to measure detection effectiveness.

In all cases, an Attack Detection Rate (ADR) was computed as round(0.05 + 100 * TruePositives /(Total Events)). The following table depicts the exercise results.

ADR [%]	Min	Max	Average
Snort well adapted	0	100	34
Snort non-Adapted	0	3.0	1.1
Esom	0	0.3	0.1
Esom adapted	n/a	n/a	n/a
Tranalyzer	0	1.1	0.4
Tranalyzer adapted	0	16.4	6.1

Fig 9 – Detection Ratio presented for Tranalyzer, ESOM and Snort.

The relatively high ratio achieved by Snort resulted from most of the offensives carried out by Red Teams being well-known attacks, thus explicitly registered by Snort (i.e. whose signature database facilitated efficiently spotting them). Interestingly, several attacks were detected and correctly labelled by Tranalyzer and ESOM while not being noticed by Snort; in most cases, this happened when attacks were little-known, used non-standard tools, leveraged unusual protocols, or exploited 0-day vulnerabilities.

Furthermore, Snort missed several covert channels episodes e.g. SSH sessions over unusual ports, or backdoor/encrypted chat or data leakage channels with random ports. ESOM on the other hand performed extremely well on identifying these events, although it was not trained at full capacity (see Problems Encountered). CDX12 thus confirmed that even under harsh training conditions RUAG's software does its job as an anomaly detection engine in encrypted environments.

The ADR distribution over the two CDX12 exercise days is depicted on the next graph. It shows that Tranalyzer, during portions of the exercise where input data was of higher quality, displayed better overall results (in addition to being more effective throughout the whole exercise at spotting unknown/new threats).

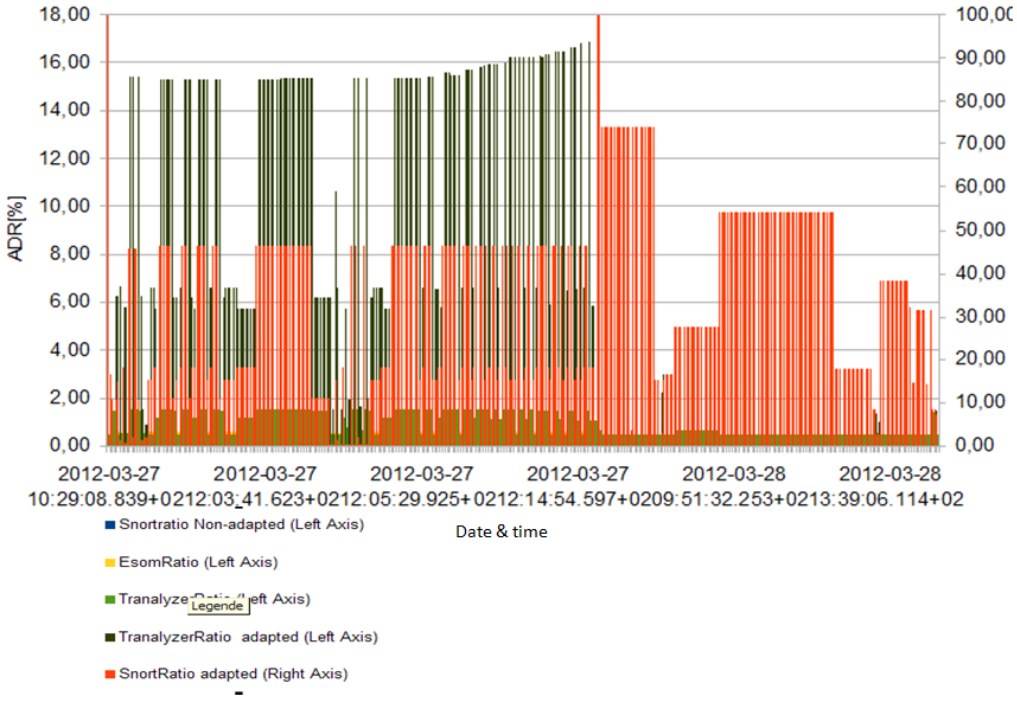

Fig 10 – ADR distribution during the two CDX12 exercise days.

Finally, the following table presents the total amount of alerts generated by each piece of software by the end of the exercise.

Tool	Number of Alerts
Snort	1’657’404
Tranalyzer	1’010’826
Esom	6’078’936

Fig 11 – Number of alerts generated by each tool.

Problems Encountered

Because RUAG's platform was running in an experimental configuration during the exercise, special attention was paid to possible problems that might arise. As partly expected, several issues that prevented RUAG's software from delivering its value at full capacity were encountered; analysis of these problems provided insight into new issues and allowed improvement to be made. The issues that took place were as follows:

- The traffic data initially obtained to train the software was acquired from Virtual Machines, which introduced noise up to 200ms into the packet inter-distance (used as a classification feature by the AI). This impaired the ESOM performance, although the maps were still able to sort IP flows into different categories with astonishing precision and recall.

- To mitigate the effect of that issue, the experiment was repeated with data acquired by the Swiss Armed Forces (FUB) – which contained less inter-distance noise, but was unsynchronised. Furthermore, relevant data for Blue Team 1 and 7 was discarded because of the crash of underlying Virtual Machines; because of this, major attack operations during this time, and for these two teams, were missed.

- Because of that crash, the software's operators were not always able to correlate the attacks documented by Red Teams with datasets at hand. This resulted in an artificial decrease in the detection rate, although the issue was partly mitigated by painstakingly comparing by hand the attack database with a list of generated timestamps.

- Several components of Tranalyzer (the application layer and the supervised KI plugins) were not used due to performance reasons in the exercise environment; therefore, several obfuscation methods (e.g. encoding, escaping) evaded the regular expression engine. This again artificially decreased the detection rate.

- The supervised ESOM plug-ins, which were not yet available in March 2012, have consequently not been used during the exercise. It is planned in the near-term to resubmit the CDX12 captured data to the high-performance Tranalyzer3, which will allow comparison of the supervised methodology with the unsupervised one.

Conclusions

The first and foremost conclusion of the CDX12 exercise is that the exercise's initial postulate is verified; namely, that Situational Awareness and theatre intelligence must be an essential keystone to any credible cyber defence operation. Without it, plans cannot be established, defences cannot be organized, offensives cannot be targeted, and damage cannot be assessed.

Furthermore, it proved that human beings alone, acting without the support of intelligent software, cannot possibly process or interpret the massive amounts of data events that occur in real-world settings. Systemic visibility is essential, but that means billions of events that only advanced technological innovation can make sense of. This is especially the case when it comes to unknown or original attack vectors.

It is only through the usage of an advanced data interpretation capacity that human beings can achieve credible Situational Awareness, and leverage that understanding to make better-informed decisions.

Empowered with such intelligence, stakeholders in the decision making process can typically leverage any of the many SIM (Security Information Management) or communication, information sharing and reporting platforms available, in order to better distribute roles and document actions.

The exercise has also revealed that Tranalyzer is an extremely powerful tool, but one that requires good expertise and training to operate properly. Tranalyzer provides a situational overview, and an awareness, that tools like Snort are utterly incapable of. Such an overview is indispensable before drilling down further on any particular issue. Future improvements in RUAG's software would nevertheless have to include a simpler setup script, out-of-the-box configurations, and a friendly User Interface.

CDX12 also showed that RUAG's ESOMs, although still experimental and not to be used by untrained analysts, can be leveraged as a uniquely innovative, highly-efficient, anomaly detection engine – which works even in encrypted environments. A possible improvement would be for the software to present stable map viewing with contextual information, thus making it easier for human operators to use.

Finally, the alarm-based view displayed at RUAG's War Room in Bern presented interesting results, and represents a promising step, especially toward detecting covert events. However, it needs to be complemented with a comprehensive log database to facilitate alarm investigation and incidents forensics.

Overall, the CDX12 exercise has proved the effectiveness of RUAG's technology, and the relevance of the company's cooperation with the Swiss Armed Forces. Despite issues encountered with the quality of source data, Tranalyzer proved its relevance and effectiveness at enhancing cyber-defence operations by enabling optimal decision-making.